



Stepwise Acquisition of Pyrimethamine Resistance in the Malaria Parasite

Citation

Lozovsky, Elena R., Thanat Chookajorn, Kyle M. Brown, Mallika Imwong, Philip J. Shaw, Sumalee Kamchonwongpaisan, Daniel E. Neafsey, Daniel M. Weinreich, and Daniel L. Hartl. 2009. Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proceedings of the National Academy of Sciences* 106(29): 12025-12030.

Published Version

doi:10.1073/pnas.0905922106

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10403692>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Classification: Biological Sciences: Evolution

Stepwise Acquisition of Pyrimethamine Resistance in the Malaria Parasite

Elena R. Lozovsky^{1*}, Thanat Chookajorn^{2*}, Kyle M. Brown¹, Mallika Imwong³, Philip J. Shaw⁴, Sumalee Kamchonwongpaisan⁴, Daniel E. Neafsey⁵, Daniel M. Weinreich⁶, Daniel L. Hartl^{1†}

¹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138. ²Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand. ³Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand. ⁴National Center for Genetic Engineering and Biotechnology (BIOTEC), Pathumthani 12120, Thailand. ⁵Broad Institute, 7 Cambridge Center, Cambridge, MA 02142. ⁶Department of Biology, Brown University, Providence, RI 02912

*These authors contributed equally to the research.

† Corresponding author. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, TEL: 617-496-3917, FAX: 617-296-5854, dhartl@oeb.harvard.edu

Manuscript information: Research report, text 21 pages, 1 table, 3 figures

Contributions: EL, TC, DW, DH designed experiments; EL, TK, PS carried out experiments; EL, TC, KB, MI, SK, DN, DW, DH analyzed data; EL, TC, KB, MI, SK, DN, DW, DH wrote paper.

Conflicts of interest: The authors have no competing or conflicting interests.

Abstract

The spread of high-level pyrimethamine resistance in Africa threatens to curtail the therapeutic lifetime of antifolate antimalarials. We have studied the possible evolutionary pathways in the evolution of pyrimethamine resistance using an approach in which all possible mutational intermediates were created by site-directed mutagenesis and assayed for their level of drug resistance. The coding sequence for dihydrofolate reductase (DHFR) from the malaria parasite *Plasmodium falciparum* was mutagenized, and tests carried out in *Escherichia coli* under conditions in which the endogenous bacterial enzyme was selectively inhibited. We studied four key amino acid replacements implicated in pyrimethamine resistance, namely N51I, C59R, S108N, and I164L. Using empirical estimates of the mutational spectrum in *P. falciparum*, and probabilities of fixation based on the relative levels of resistance, the predicted favored pathways of drug resistance are consistent with previous kinetic studies as well as DHFR polymorphisms observed in natural populations. We find that three pathways account for nearly 90% of the simulated realizations of the evolution of pyrimethamine resistance. The most frequent pathway (S108N and then C59R, followed by N51I and then I164L) accounts for more than half of the simulated realizations. Our results also suggest an explanation of why I164L is detected in Southeast Asia and South America, but not so far at significant frequencies in Africa.

\body

Understanding the molecular evolution of drug resistance has potential clinical implications for the development of therapeutic protocols to forestall resistance, the rational design of modified drugs that target resistant proteins, and the deployment of more effective drugs less likely to promote resistance. More generally, understanding the acquisition of drug resistance can reveal broader features of the multidimensional fitness landscape that organisms traverse as they evolve, and the constraints that this landscape imposes on their evolution.

The malaria parasite offers promising opportunities for such studies, as the useful therapeutic lifetime of most first-line malaria drugs has been compromised by the evolution of parasite resistance. Notable examples include chloroquine (1), atovaquone (2), and pyrimethamine (3, 4). Both chloroquine resistance and pyrimethamine resistance are associated with multiple amino acid replacements in the target protein (5, 6). Mutations in other parasite genes may also contribute quantitatively to increase the resistance phenotype or enhance parasite fitness (7-9).

In this study we focus on the evolution of resistance to pyrimethamine, a drug that competitively inhibits the parasite enzyme dihydrofolate reductase [DHFR, E.C. 1.5.1.3]. DHFR is part of a bifunctional enzyme needed for the synthesis of tetrahydrofolate, an essential precursor of purines and several amino acids (10, 11). While pyrimethamine and other antifolates are currently the first-line treatment of malaria in many parts of Africa (12-14), high-level resistance has been reported elsewhere (4, 13, 15). The genetic basis of pyrimethamine resistance in *P. falciparum* is associated with a small number of amino acid replacements in the parasite DHFR (3, 4, 16-18). Each of these replacements has been observed in clinical isolates of resistant strains (14, 19-21). Resistant mutants in DHFR are generally associated with some

combination of four amino acid changes—in particular, Asn51 to Ile (N51I), Cys59 to Arg (C59R), Ser108 to Asn (S108N), and Ile164 to Leu (I164L). These four replacements affect the enzyme-binding pocket and reduce the binding affinity for pyrimethamine (22). Some of the mutants also compromise the endogenous DHFR enzymatic efficiency (6). Resistant mutants of DHFR appear to have arisen in Southeast Asia and then spread to Africa (23). No other point mutations in the DHFR gene are commonly associated with drug resistance, although a few rare variants associated with heightened resistance have been reported (24).

In view of the continuing clinical importance of antifolates in treating malaria in Africa, the pathways by which resistance may have evolved are of immediate interest. While protein evolution proceeds largely via the sequential replacement of individual amino acids, there are processes, trade-offs, and interactions that can constrain the temporal order in which these replacements occur. Factors that may constrain the realized evolutionary pathways or trajectories include:

- Mutation bias, which in the malaria parasite is a strong bias toward AT, and therefore favors AT-rich codons over others (25);
- Adaptive conflict, because some amino acid replacements that increase drug resistance may also impair the endogenous function of the protein (6).
- Epistasis resulting from non-additive interactions between mutant sites in the same gene, because the effect of any amino acid replacement may depend on the sequence context in which it occurs (26).
- Epistasis resulting from non-additive interactions between mutations in different genes, because the evolution of any enzyme in a metabolic pathway alters the selection pressures acting on other enzymes in the same metabolic pathway or others (27).

In principle, the evolutionary pathways of drug resistance could be reconstructed by selection for resistance in cultured isolates. Experimental systems in which the parasite enzyme is studied in either transgenic bacteria (17) or yeast (28) have been developed, and both approaches have been successful in identifying amino acid sites important in clinical drug resistance as well as novel resistance sites not observed in nature. Kinetic studies of putative intermediates found as polymorphisms in natural populations have also proven informative in inferring the stepwise evolution of drug resistance (6). The results suggest three likely pathways by which pyrimethamine resistance may have evolved (6).

In the present experiments, we adopted a third approach, which is to genetically engineer and analyze the phenotypes of all possible combinations of the amino acid replacements implicated in drug resistance. In previous experiments carried out to examine possible pathways for the evolution of various functions in a number of proteins (29-32), epistasis among the mutant sites substantially reduced the number of selectively accessible mutational pathways. Mutational bias and adaptive conflicts were not explicitly addressed. These limitations were addressed in the present analysis in two ways. First, we assumed the mutational biases implied by changes in genome sequence in the evolutionary lineage of *P. falciparum*. Second, in the experimental studies, we used the transgenic bacterial system in which each bacterial strain contained an alternatively mutated DHFR coding sequence from *P. falciparum*. The endogenous bacterial enzyme was inhibited with an antifolate to which the parasite enzyme is naturally resistant. The transgenic mutant genotypes therefore have an adaptive conflict that trades resistance to inhibition by pyrimethamine against efficient conversion of dihydrofolate into tetrahydrofolate, which impinges on growth rate in the absence of pyrimethamine. The results suggest that this trade-off may provide a hypothesis to help

explain why the quadruple DHFR mutant so widespread in Southeast Asia has not yet spread in Africa.

Overall, our results confirm but also extend previous hypotheses on the evolutionary pathways of pyrimethamine resistance. Our inferences take earlier results a step farther in being able to estimate a probability for the realization of each possible pathway of the evolution of pyrimethamine resistance. The most frequent pathway accounts for more than half of all realizations in simulated evolutionary replays, and the top three account for 87.5% percent of the realizations. Our results are discussed in the contexts of clinical resistance of *P. falciparum* to antifolates, protein folding and function, and more generally protein evolution.

Results

By means of site-directed mutagenesis of a bacterial plasmid containing the coding sequence of *P. falciparum* DHFR, we created all 16 possible combinations of the mutants Asn51Ile (N51I), Cys59Arg (C59R), Ser108Asn (S108N), and Ile164Leu (I164L). The sequences of the mutagenized DHFR genes were confirmed by direct sequencing.

Drug-Resistance Assays

Assays for pyrimethamine resistance were carried out by means of a bacterial complementation system (17). In this system, the endogenous bacterial enzyme was chemically inhibited by 5 μ M trimethoprim, a concentration sufficient to prevent bacterial growth. The *P. falciparum* DHFR enzyme is insensitive to trimethoprim, and so can complement the inhibited bacterial enzyme. The bacterial system therefore requires that the DHFR enzyme activity be sufficient to allow the cells to grow, which in the presence of pyrimethamine represents an adaptive conflict between drug resistance and catalytic activity.

The level of pyrimethamine resistance was estimated as the concentration of drug that inhibited cell growth by 50 percent, a metric referred to as inhibitory concentration 50 and symbolized IC₅₀. Bacterial strains carrying each of the DHFR alleles were assayed across a range of pyrimethamine concentrations, and growth rate as a function of pyrimethamine concentration was estimated by polynomial regression. To determine the empirical distribution of IC₅₀ estimates, the IC₅₀ value for each of the 16 DHFR alleles was estimated in 4–8 independent biological replicates, each with 4 technical replicates.

The results are shown in Fig. 1 in the form of a bar graph, and in tabular form in the Table S1. The DHFR alleles indicated along the horizontal axis are given in the form of a vector of 0's and 1's that correspond left-to-right to residues 51, 59, 108, and 164. Each 0 indicates a nonmutant codon and each 1 a mutant codon. The symbol 0000 therefore represents the nonmutant sequence NCSI whereas 1111 represents the quadruple mutant IRNL. The error bars are equal to the standard deviation of each of the distributions of estimated IC₅₀ values. Larger values of IC₅₀ can be estimated with greater precision than smaller ones. The coefficient of variation is about 30% of the mean for IC₅₀ values < 10 µg/ml, and about 5% of the mean for values > 10 µg/ml.

Fig. 1 shows that two of the four possible single mutants (1000 and 0001) are about equally sensitive to pyrimethamine as the nonmutant allele 0000. The first step in the evolutionary pathway toward pyrimethamine resistance is therefore likely to be either 0010 or 0100. Beyond that, the favored evolutionary pathways are not clear from inspection, although the quadruple mutant is obviously much more resistant than any of the triple mutants.

While the average IC₅₀ of the single, double, and triple mutants steadily increases from 3.06 µg/ml to 49.9 µg/ml to 137.6 µg/ml, respectively, the complexity of the drug-resistance landscape is evident from Fig. 1. Compare the values among the

double mutants, for example, whose IC₅₀s range from 0 to almost 150 µg/ml. A notable example of epistatic interaction among the mutant sites is allele 0011, which is extremely sensitive to pyrimethamine, whereas its single-mutant constituents 0010 and 0001 have IC₅₀ values of 9.56 µg/ml and 0.29 µg/ml, respectively. The growth of strains containing 0011 is impaired even in the absence of pyrimethamine. In particular, the growth rate of 0011 relative to that of 0000 in the absence of the drug is 0.169 ± 0.002 , whereas the relative growth rates of 0010 and 0001 are 1.103 ± 0.009 and 0.968 ± 0.004 , respectively. In other words, the combination of two mutants, each with a virtually normal growth rate in the absence of drug, results in a double mutant whose growth is severely impaired. The inability of 0011 to grow in pyrimethamine is a good illustration of the adaptive conflict between resistance to inhibition and residual enzyme activity.

The IC₅₀ values in Fig. 1 are in good agreement with another widely used metric of drug resistance, the minimal inhibitory concentration (MIC), which refers to the smallest concentration of the drug that can completely inhibit growth (Table S1). Although the solubility of pyrimethamine limits the discrimination among the MIC values of the most highly resistant alleles, the concordance between the estimated IC₅₀ and MIC values is very high. Because it is the relative ranking of the DHFR alleles that matters most in inferring the possible evolutionary pathways, this is the relevant comparison, and the linear association exhibits $r^2 = 0.81$. The linear association between the raw resistance scores of IC₅₀ and MIC has $r^2 = 0.68$, but as noted this comparison is less relevant. The close correspondence between the rank ordering of alleles according to IC₅₀ and MIC seems to exclude the possibility that the analysis of the evolutionary pathways to pyrimethamine resistance is unduly influenced by the choice of resistance assay.

Evolutionary Model

To analyze the data in Fig. 1, we employed an evolutionary model in which selection acts to increase pyrimethamine resistance. Selection pressure is assumed to be strong relative to mutation pressure, which implies that the time to fixation or loss of a newly arising mutation is much shorter than the time between the occurrence of new mutations. We also assume that the population size is sufficiently large that random genetic drift has a negligible effect on the probability of fixation. In this model, the selectively driven fixation or loss of each newly arising mutant takes place before the next mutation occurs, and so an evolutionary trajectory consists of a succession of new mutant alleles becoming fixed, each of which increases the level of drug resistance (30). In the mutational process, we allow all possible single-mutant neighbors to occur, including reversions of previously fixed mutations (33). Selection was based only on the relative values of IC₅₀, and differences in growth rate in the absence of drug were not explicitly taken into account. The rationale is that IC₅₀ values differ among alleles by a factor of more than 1000, whereas for most alleles the growth rates in the absence of drug differ by no more than a factor of 2 (Table S2). The presence of any significant drug pressure therefore swamps selection attributable to intrinsic differences in growth rate; however, as discussed later, the case of the quadruple mutant is especially interesting in the light of potential compensatory mutations.

The evolutionary model also incorporated the mutational bias of the malaria parasite. The genome of *P. falciparum* consists of about 82% A–T nucleotide pairs (34), reflecting a strong mutational bias toward A or T (25). The mutation matrix (Table S3) was estimated using a genome-wide collection of 1105 single-nucleotide polymorphisms (SNPs) in intergenic regions, culled from publicly available sequencing reads from *P. falciparum* and its sister species *P. reichenowi*. Methods of analysis are explained in the text accompanying Table S3.

Analysis of Evolutionary Pathways

We used computer simulation to explore the evolutionary implications of the results in Fig. 1. A flow chart of the computer algorithm is provided in Fig. S1. Briefly, replicate evolutionary landscapes were defined by random sampling of IC50 values for each of the 16 possible alleles from normal distributions with the allele-specific genotypic means and standard deviations depicted in Fig. 1. The possible evolutionary paths along each landscape were then explored by randomly choosing single-step mutations, including possible reverse mutations, according to the mutation model in Table S3. Each new mutant allele was discarded if its IC50 value was smaller than that of the prevailing allele, and otherwise the mutant was fixed or lost according to the difference between the ranks of the IC50 values defining the evolutionary landscape. This procedure reflects the principle that the probability of fixation of a new mutation is proportional to its selective advantage, while the use of ranks renders the multitude of simulated fitness landscapes commensurate. Evolution on each simulated landscape was continued until either fixation of the allele with the maximum IC50 occurred, or else fixation of a different allele at a submaximal fitness peak took place, whose single-mutant neighbors all had lower IC50s.

Simulation of 10,000 landscapes with 1000 independent excursions in each landscape identified 10 evolutionary pathways that were traversed at significant, though very uneven, frequencies. These evolutionary pathways are shown in red in Fig. 2, where the vectors of 0's and 1's again depict the amino acid residues 51, 59, 108, and 164. A large number of other pathways were realized, each at a negligible frequency, which included all pathways with mutational reversions. These rare pathways mainly reflected the random sampling of extreme values from the distributions in Fig. 1, resulting from the large number of simulated landscapes.

To estimate the relative probabilities of the prominent pathways more precisely, we chose 100 of the landscapes at random and explored each with one million independent evolutionary excursions. The results confirmed the prominence of the 10 previously mentioned pathways, and yielded the estimates of the relative probability of each pathway shown at the right in Fig. 2 in the column headed "Prob." Only those evolutionary pathways traced in red constitute the top 10 in the simulations, with the more likely pathways traced in thicker lines.

The relative frequencies of the 10 major pathways were virtually identical when 10,000 landscapes were traversed one thousand times each, or when 100 landscapes were traversed one million times each. This consistency reflects the relatively small standard deviations among the IC50 distributions (Fig. 1). The probabilities in Fig. 2 are primarily determined by the differences in IC50, and are only slightly affected by the magnitude of the mutation bias.

Just three pathways in Fig. 2 account for 87.5% of those traversed at significant frequencies in the simulations. All three of these pathways feature S108N as the first step, and the two most likely include C59R as the second step. Similarly, all three of the most likely pathways have I164L incorporated either last or else next to last.

Comparison with DHFR Polymorphisms

The results in Fig. 2 are consistent with the combinations of amino acid replacements observed at significant frequencies in worldwide surveys of *P. falciparum* (Table 1, Refs. in Table S4). The common polymorphisms are indicated by the red ovals in Fig. 2, and they coincide with the intermediates predicted in the most likely pathways.

Note in Table 1 that the quadruple mutant has not yet been found at significant frequencies in Africa. The absence of this allele in Africa is all the more surprising in light of the compelling evidence that pyrimethamine resistant alleles in Africa derived

from those present in Southeast Asia, where the quadruple mutant is quite common (23). Our results may also bear on this issue. In the absence of pyrimethamine, the growth rates of the strains bearing 0010 (S108N), 0110 (S108N + C59R), or either of the triple mutants, is not manifestly impaired relative to the previous step or even the nonmutant allele (Fig. 3, Table S2). These results are consistent with previous inferences based on comparisons of the Michaelis constant (K_m) and catalytic turnover rate (k_{cat}) of the wildtype and mutant enzymes (35). However, when the fourth mutation is added, either N51I or I164L, creating the quadruple mutant, there is a strongly deleterious interaction between N51I and I164L, with the result that the fourth mutation decreases growth rate in absence of pyrimethamine by 30–40% (Fig. S2), even though it increases IC₅₀ by 25–50% (Fig. 1). Hence, in the absence of compensatory mutation (see Discussion), the potential fitness cost of incorporating the last mutation in the pathway is substantial.

Discussion

Previous *in vitro* studies of the mutant DHFR proteins and their susceptibility to inhibition by pyrimethamine have identified three hypothetical pathways for the stepwise evolution of resistance (6). These are precisely the three most likely pathways identified in Fig. 2. Our data take the analysis a step farther in providing estimates of the relative likelihoods of the pathways. The order of amino acid replacements S108N, C59R, N51I, I164L is favored by a factor of about two over S108N, C59R, I164L, N51I, and the latter is favored by another factor of about two over S108N, N51I, C59R, I164L. These three pathways together account for 87.5% of the excursions over the evolutionary landscape that occur with significant frequencies based on the data in Fig. 1.

Pyrimethamine acts by competing with dihydrofolate for access to the binding pocket of DHFR. Because endogenous DHFR activity is essential for viability, the evolution of resistance occurs through increased substrate specificity (36). The key kinetic parameters are the Michaelis constant (K_m), the pyrimethamine dissociation constant (K_i), and the catalytic turnover rate (k_{cat}) (6). The competitive binding dynamics are biphasic, which implies that the rate of product production is largely independent of catalytic capacity (k_{cat}/K_m) if the relative preference of the enzyme for substrate (K_i/K_m) is below some threshold. Consistent with this picture, the pathways with the highest probabilities in Fig. 2 all improve the substrate specificity at the expense of catalytic capacity. According to Table 3 in Ref. (6), relative to the nonmutant DHFR, the quadruple mutant shows an approximate 500-fold increase in K_i/K_m at the expense of about a 6-fold decrease in k_{cat}/K_m . Interestingly, among the single mutants, S108N yields the largest increase in IC50 relative to the nonmutant allele, and it is clinically the most important (14, 21), even though C59R yields both a larger increase in both K_i/K_m and k_{cat}/K_m (6). This observation implies that protein attributes beyond kinetic constants alone also contribute to differences in IC50. Obvious candidates include protein folding, stability, potential for aggregation, rate of degradation, and so forth, and the role of natural selection acting through such factors has been discussed (37). While DHFR is an established model system for the study of protein folding (38), such data for the pyrimethamine-resistant DHFR alleles are as yet unavailable.

The fitness cost of the quadruple mutation in the absence of pyrimethamine (Fig. 3) may help explain why the quadruple mutant has not yet spread in Africa. An intriguing link between the evolution of drug resistance and the intensity of malaria transmission has been noted: Drug resistance may evolve more readily in areas of low transmission (7). A possible mechanism is that some drug-resistance mutants may have deleterious effects on fitness in the absence of the drug, which may be tolerated because of compensatory mutations that occur elsewhere in the genome (8). In areas of low

transmission, inbreeding is relatively high and genetic recombination restricted, and so the deleterious drug-resistance determinant and the compensatory mutations can remain genetically associated. In areas of high transmission, however, frequent mixed infectious result in genetic recombination, which breaks down the association between the mutant genes.

Consistent with the hypothesis of compensatory mutation, a copy-number polymorphism in the gene encoding GTP-cyclohydrolase I, the first gene in the folate biosynthetic pathway, has recently been shown to be associated with the DHFR quadruple mutant in Thailand (9). Elsewhere in the genome, a selective sweep associated with pyrimethamine resistance has been reported across a region of chromosome 13 (39). It is an interesting, testable hypothesis that this region of chromosome 13 may include one or more mutations that also alleviate deleterious fitness effects of DHFR resistance alleles.

While previous studies of mutationally accessible pathways to high-fitness sequences (29-32, 40) have explored the evolution of novel functions, pathways were not penalized if they simultaneously caused the organism to lose other functionality. In this sense, our experiment provides the first glimpse of the constraint imposed on evolutionary pathways by the adaptive conflict between inhibitor resistance and maintenance of sufficient endogenous activity. Nevertheless, levels of constraint on pathways in DHFR are comparable to previous work, in which constraint was a reflection of epistasis in a single trait, and many but not all evolutionary pathways are rendered inaccessible to selection.

Of great interest in exploring evolutionary landscapes is the possible existence of submaximal fitness peaks, at any of which a population may become stranded because every one-step mutation has a lower fitness. The DHFR landscape features one such submaximal peak. The sequence 1001 (N51I + I164L) has an IC₅₀ of about 100 µg/ml,

and it is surrounded on the landscape by 1101 (N51I + C59R + I164L), 1011 (N51I + S108N + I164L), 1000 (N51I), and 0001 (I164L), which have IC₅₀s of about 56, 57, 0.4, and 0.3 µg/ml, respectively (Fig. 1). Hence 1001 is accessible through either 1000 or 0001, but neither pathway is realized frequently enough to be among the top 10 (Fig. 2), mainly because among the single mutants, the mutation yielding S108N (G–C to A–T) is strongly favored to be fixed. Interestingly, the low fitness of the quadruple mutant in the absence of pyrimethamine (Fig. 3) is not observed in 1001 (N51I + I164L). Evidently, the deleterious interaction between N51I and I164L occurs only on a background of C59R + S108N.

The observed polymorphisms in natural populations of *P. falciparum* in Table 1 are consistent with the three major pathways identified in our analysis. Such agreement would not necessarily be expected. Drug treatment in a laboratory setting is carefully controlled and reproducible, whereas that in the field is variable depending on drug dosage, frequency, potency, pharmacokinetics, compliance, and other factors. Moreover, our results were obtained from studies in the *E. coli* system, not in *P. falciparum* itself, and some significant disparities might have been anticipated in view of the vast evolutionary distance between these organisms. It is therefore reassuring that the inferred major pathways receive additional support from the naturally occurring polymorphisms as well as from the kinetic studies discussed earlier. More generally, such support augers well for a similar use of model organisms in future studies of the evolution of drug resistance, not only in the malaria parasite, but also in many other organisms.

Methods

The sequence coding for DHFR amino acids 1–239 was isolated from *P. falciparum* strain K1 by the polymerase chain reaction and cloned into the vector pET17 (Novagen) without the

T7 tag. The resulting plasmid was transformed into *E. coli* strain HMS174(DE3) (Novagen), and the construct confirmed by sequencing. DHFR mutants were introduced by QuikChange site-directed mutagenesis (Stratagene) and reintroduced into strain HMS174(DE3). Residues Asn51, Cys59, Ser108 and Ile164 were changed to Ile, Arg, Asn, and Leu, respectively, individually and in all possible combinations, using oligonucleotides (Operon). The mutants were all confirmed by DNA sequencing.

Estimates of IC₅₀ (Table S1) were obtained as follows. Overnight cultures grown in rich LB medium, containing 50 ug/ml of carbamicillin to maintain the DHFR plasmid, were diluted and the cells allowed to grow into log phase. Aliquots of these cultures were diluted to approximately 10⁷ cells/ml in a series of concentrations of pyrimethamine in LB containing 5 uM of trimethoprim, and dispensed into the wells of 96-well microtiter plates. The plates were incubated in the dark at 37°C for 19–20 hours, and the resulting cell concentrations estimated spectrophotometrically from the optical density at 600 nm. Preliminary experiments identified the interval in which the IC₅₀ was likely to be found, and final estimates were based on concentrations bracketing this interval. Growth rate as a function of pyrimethamine concentration was estimated by polynomial regression, and the IC₅₀ value estimated from this curve. To minimize the experimental error as much as possible, the overall IC₅₀ for each of the 16 DHFR alleles was estimated from the IC₅₀ values observed in 4–8 independent biological replicates and 4 technical replicates of each biological replicate.

For the MIC assays (Table S1), bacterial strains were streaked onto fresh agar medium consisting of LB broth and trimethoprim plus various concentrations of pyrimethamine in increments of 25 µg/ml, and the plates were incubated at 37°C for 18 hours. For each DHFR allele, the MIC was estimated with a minimum of three biological replicates.

Excursions through evolutionary landscapes were simulated by means of an algorithm implemented in PERL. A flow chart of the algorithm is shown in Fig. S1. Mean pathway probabilities and their 95% confidence intervals were estimated using scripts written in R (v2.2.1, The R Foundation for Statistical Computing). Ad hoc analysis and computer graphics were carried out in Mathematica (Wolfram Research).

Acknowledgments

We are grateful to Sarah Volkman, Yongyuth Yuthavong, and Prapon Wilairat for their thoughtful advice, and David Houle and Michael Ferdig for their review and suggestions regarding the manuscript. This work was supported by NIH grant R01GM079536 to DLH. TC is supported by grant MRG5080418 from the Thailand Research Fund, Commission on Higher Education.

References

1. Welles TE (2002) *Plasmodium* chloroquine resistance and the search for a replacement antimalarial drug. *Science* 198:124-126.
2. Looareesuwan S, *et al.* (1996) Clinical studies of atovaquone, alone or in combination with other antimalarial drugs, for treatment of acute uncomplicated malaria in Thailand. *Am J Trop Med Hyg* 54:62-66.
3. Peterson DS, Walliker D, Welles TE (1988) Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. *Proc Natl Acad Sci U S A* 85:9114-9118.
4. Snewin V, A., England SM, Sims PFG, Hyde JE (1989) Characterisation of the dihydrofolate reductase-thymidylate synthetase gene from human malaria parasites highly resistant to pyrimethamine. *Gene* 76:41-52.

5. Fidock DA, *et al.* (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell* 6:861-871.
6. Sirawaraporn W, Sathitkul T, Sirawaraporn R, Yuthavong Y, Santi DV (1997) Antifolate-resistant mutants of *Plasmodium falciparum* dihydrofolate reductase. *Proc Natl Acad Sci U S A* 94:1124-1129.
7. Hastings IM (2003) Malaria control and the evolution of drug resistance: an intriguing link. *Trends Parasitol* 19:70–73.
8. Nzila A, *et al.* (2006) Why has the dihydrofolate reductase 164 mutation not consistently been found in Africa yet? *Trans R Soc Trop Med Hyg* (2005) 99, 99:341—346.
9. Nair S, *et al.* (2008) Adaptive copy number evolution in malaria parasites. *PLoS Genet* 4:e1000243.
10. Patel OG, Mberu EK, Nzila AM, Macreadie IG (2004) Sulfa drugs strike more than once. *Trends Parasitol* 20:1-3.
11. Kompis IM, Islam K, Then RL (2005) DNA and RNA synthesis: Antifolates. *Chem Rev* 105:593-620.
12. Sibley CH, *et al.* (2001) Pyrimethamine-sulfadoxine resistance in *Plasmodium falciparum*: What next? *Trends Parasitol* 17:582-588.
13. Wongsrichanalai C, Pickard AL, Wernsdorfer WH, Meshnick SR (2002) Epidemiology of drug-resistant malaria. *Lancet Infect Dis* 2:209-218.
14. Kublin JG, *et al.* (2002) Molecular markers for failure of sulfadoxine-pyrimethamine and chlorproguanil-dapsone treatment of *Plasmodium falciparum* malaria. *J Infect Dis* 185:380-388.
15. Roper C, *et al.* (2003) Antifolate antimalarial resistance in southeast Africa: A population-based analysis. *Lancet* 361:1174-1181.
16. Ferlan JT, Mookherjee S, Okezie IN, Fulgence L, Sibley CH (2001) Mutagenesis of dihydrofolate reductase from *Plasmodium falciparum*: Analysis in *Saccharomyces cerevisiae*

- of triple mutant alleles resistant to pyrimethamine and WR99210. *Mol Biochem Parasitol* 113:139-150.
17. Chusacultanachai S, Thiensathit P, Tarnchompoo B, Sirawaraporn W, Yuthavong Y (2002) Novel antifolate resistant mutations of *Plasmodium falciparum* dihydrofolate reductase selected in *Escherichia coli*. *Mol Biochem Parasitol* 120:61–72.
18. Plowe CV (2009) The evolution of drug-resistant malaria. *Trans Roy Soc Trop Med Hyg* 103, Suppl. 1:S11-S14.
19. Wang P, et al. (1997) Resistance to antifolates in *Plasmodium falciparum* monitored by sequence analysis of dihydropteroate synthetase and dihydrofolate reductase alleles in a large number of field samples of diverse origins. *Mol Biochem Parasitol* 89:161-177.
20. Sims PFG, Wang P, Hyde JE (1999) On the efficacy of antifolate antimalarial combinations in Africa. *Parasitol Today* 15:132-134.
21. Nzila AM, et al. (2000) Toward an understanding of the mechanism of pyrimethamine-sulfadoxine resistance in *Plasmodium falciparum*: Genotyping of dihydrofolate reductase and dihydropteroate synthase in Kenyan parasites. *Antimicrob Agents Chemother* 44:991-996.
22. Yuthavong Y, et al. (2005) Malarial (*Plasmodium falciparum*) dihydrofolate reductase-thymidylate synthase: Structural basis for antifolate resistance and development of effective inhibitors. *Parasitol* 130: 249-259.
23. Roper C, et al. (2004) Intercontinental spread of pyrimethamine-resistant malaria. *Science* 305:1124.
24. Bates SJ, et al. (2004) Rare, highly pyrimethamine-resistant alleles of the *Plasmodium falciparum* dihydrofolate reductase gene from 5 Africa Sites. *J Infect Dis* 190:1783-1792.
25. DePristo MA, Zilversmit MM, Hartl DL (2006) On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene* 378:19-30.

26. Weinreich DM,Chao L (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* 59:1175-1182.
27. Keightley PD,Kacser H (1987) Dominance, pleiotropy and metabolic structure. *Genetics* 117:319-329.
28. Sibley CH,Macreadie IG (2001) Novel approaches to tackling malarial drug resistance using yeast. *IUBMB Life* 52:285-289.
29. Lunzer M, Miller SP, Felsheim R,Dean AM (2005) The biochemical architecture of an ancient adaptive landscape. *Science* 310: 499-501.
30. Weinreich DM, Delaney NF, DePristo MA,Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312:111-114.
31. Reetz MT, Wang L-W,Bocola M (2006) Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space. *Angewadte Chemie* 118:1258-1263.
32. Bridgham JT, Carroll SM,Thornton JW (2007) Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312:97-100.
33. DePristo MA, Hartl DL,Weinreich DM (2007) Mutational reversions during adaptive protein evolution. *Mol Biol Evol* 24:1608-1610.
34. Gardner MJ, *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419:498-511.
35. Sandefur CI, Wooden JM, Quaye IK, Sirawaraporn W,Sibley CH (2007) Pyrimethamine-resistant dihydrofolate reductase enzymes of *Plasmodium falciparum* are not enzymatically compromised *in vitro*. *Mol Biochem Parasitol* 154:1-5.
36. Rastelli G, *et al.* (2000) Interaction of pyrimethamine, cycloguanil, WR99210 and their analogues with *Plasmodium falciparum* dihydrofolate reductase: Structural basis of antifolate resistance. *Bioorganic and Medicinal Chemistry* 8:1117-1128.

37. DePristo MA, Weinreich DM, Hartl DL (2005) Missense meandering through sequence space. *Nature Rev Genet* 6:678-687.
38. Wang Z, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *Journal of Molecular Biology* 320:85-95.
39. Volkman SK, et al. (2007) A genome-wide map of diversity in *Plasmodium falciparum*. *Nature Genet* 39:113-119.
40. Poelwijk FJ, Kiviet DJ, Sans SJ (2006) Evolutionary potential of a duplicated repressor-operator pair: Simulating pathways using mutation *PLoS Comput Biol* 2:e58.

Figure Legends

Fig. 1. Mean IC₅₀ values for pyrimethamine among the 16 possible combinations of mutant amino acid sites in DHFR. The error bars are the standard deviations of the estimated sampling distributions of IC₅₀.

Fig. 2. Major inferred pathways for the evolution of pyrimethamine resistance. The top 10 pathways are shown in red along with their estimated probabilities.

Fig. 3. Relative growth rates of each intermediate in the two major evolutionary pathways. The growth rate of each strain after each step in the pathway is measured relative to strain carrying the immediately preceding allele. The final steps in the pathway 0110–1110–1111 are shown in orange, those in the pathway 0110–0111–1111 are shown in red. The error bars are 95% confidence intervals.

Table legends

Table 1. Common polymorphic DHFR alleles in *P. falciparum*